AFRL-RH-WP-TR-2012-0170

Automated Analysis and Classification of Infected Macrophages Using Bright-Field Amplitude Contrast Data

Umesh Adiga
Brian Bell
UES, Inc.
4401 Dayton Xenia Road
Beavercreek, OH 45432

Larissa Ponomareva
Sandra Nelson
University of Cincinnati
Drug Discovery Center
2180 E. Galbraith Road
Cincinnati, OH 45237

Stephen Kanzleman
Henry M. Jackson Foundation
6720-A Rockledge Drive
Bethesda, MD 20817

Debbie Taylor
Ryan Kramer
Thomas Lamkin
Human-Centered ISR Division
Human Signatures Branch

August 2012
Interim Report

AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING, HUMAN
EFFECTIVENESS DIRECTORATE, WRIGHT-
PATTERSON AIR FORCE BASE, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

//Signed//   //Signed//

_____   _____
Thomas Lamkin, PhD.   Louise A. Carter, PhD.
Work Unit Manager   Chief, Human-Centered ISR Division
Human Signatures Branch   Human Effectiveness Directorate
                                                     711th Human Performance Wing
                                                     Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
30 08 12	Interim	03/01/2010 – 03/31/2011

4. TITLE AND SUBTITLE

Automated Analysis and Classification of Infected Macrophages using Bright-Field Amplitude Contrast Data

5a. CONTRACT NUMBER
IN-HOUSE

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Umesh Adiga[1] Debbie Taylor[2]
Brian Bell[1] Ryan Kramer[2,]
Larissa Ponomareva[3] Thomas J. Lamkin [2]
Sandra Nelson[3]
Stephen Kanzlemar[4]

5d. PROJECT NUMBER
ODT

5e. TASK NUMBER
WP

5f. WORK UNIT NUMBER
H066 (ODTWP002)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

[1]UES Inc.
4401 Dayton Xenia Rd
Beavercreek, OH 45432

[3]University of Cincinnati
Drug Discovery Center
2180 Galbraith Road
Cincinnati, OH 45237

[4]Henry M. Jackson
6720-A Rockledge Dr.
Bethesda, MD 20817

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

[2]Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Human-Centered ISR Division
Human Signatures Branch
Wright-Patterson Air Force Base, OH 45433

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
711 HPW/RHXB

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RH-WP-TR-2012-0170

12. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution A. Approved for public release; distribution unlimited; 88ABW-2011-2636, cleared 10 May 11

13. SUPPLEMENTARY NOTES

Report contains color.

14. ABSTRACT

In this paper we demonstrate that bright-field amplitude contrast image data can be used in high throughput screening (HTS) for simultaneous measurement of cell density, cell viability and general classification of individual cells into phenotypic classes such as infected or uninfected without the need to use fluorescent dyes. We present a robust image analysis pipeline where the original data is subjected to image standardization, noise reduction and, image enhancement filters and segmentation by region growing. We further show that by implementing a robust set of image enhancement and data standardization filters, it is possible to use bright-field amplitude contrast data to count and analyze the cells without requiring specimen fixation and/or use of fluorescent dyes. This work develops new, faster and less expensive, reliable imaging techniques for live cell analysis in HTS and successfully addresses a particular need for direct measurement of cell density in HTS.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	SAR	38	Thomas J. Lamkin
Unclassified	Unclassified	Unclassified			**19b. TELEPHONE NUMBER** *(Include Area Code)* N/A

. . . .

THIS PAGE IS INTENTIONALLY LEFT BLANK.

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE IS INTENTIONALLY LEFT BLANK.

Automated Analysis and Classification of Infected Macrophages using Bright-field Amplitude Contrast Data

Umesh Adiga[1], Debbie Taylor[2,5], Brian Bell[1], Larissa Ponomareva[3], Stephen Kanzlemar[4], Ryan Kramer[2,5], Sandra Nelson[3], Thomas J. Lamkin[2,5,6]

[1]UES Inc.

[2]Air Force Research Laboratory, 711th HPW/RHPCB, Wright Patterson Air Force Base, OH

[3]The University of Cincinnati, Drug Discovery center, Metabolic Disease Institute, Cincinnati, OH

[4]Henry M. Jackson Foundation

[5]The University of Cincinnati, Dept of Molecular Genetics, Biochemistry and Microbiology, Cincinnati, OH

[6] Corresponding author: Thomas J. Lamkin, Air Force Research Laboratory, 711th HPW/RHPCB; thomas.lamkin@wpafb.af.mil

Abstract: In this paper we demonstrate that bright-field amplitude contrast image data can be used in high throughput screening (HTS) for simultaneous measurement of cell density, cell viability and general classification of individual cells into phenotypic classes such as infected or uninfected without the need to use fluorescent dyes. We present a robust image analysis pipeline where the original data is subjected to image standardization, noise reduction and,

image enhancement filters and segmentation by region growing. We further show that by implementing a robust set of image enhancement and data standardization filters, it is possible to use bright-field amplitude contrast data to count and analyze the cells without requiring specimen fixation and/or use of fluorescent dyes. This work develops new, faster and less expensive, reliable imaging techniques for live cell analysis in HTS and successfully addresses a particular need for direct measurement of cell density in HTS.

Introduction

High content screening (HCS) has evolved into a mature technique that typically uses automated fluorescence microscopy of cells stained with multiple fluorescent labels. A general approach to HCS of compound toxicity, e.g. - drug-discovery, is to use different dyes that co-localize with protein or cell organelles to evaluate the effect of compounds on those cellular components. In case of infection studies one may use engineered self‑fluorescing bacteria or fluorescent tags to bacterial proteins to determine multiplicity of infection (MOI) in addition to fluorescent dyes to evaluate the host cell phenotype. Unfortunately, fluorescent dyes can have undesirable toxic effects, produce variations in fluorescent intensity, or result in spectral overlap that adversely affects resolution of image-based phenotypic studies. Bright field imaging is a non-toxic, dye free technique to image cells that measures direct light transmission through the cell.

2

Bright field microscopy utilizes direct observation of the illuminated objects and is widely used to observe cells under the microscope. A number of studies have been published on the usefulness of bright field imaging in cell detection and analysis. While several researchers (1, 2, 3), have used phase contrast microscopy to visualize the phase shift introduced by the interaction of light with the objects, these techniques require special objectives and devices. On the other hand, bright field amplitude contrast microscopy does not require special devices although the image contrast is quite poor owing to the transparency of the cells.

Bright field images in the current study are the result of overall amplitude contrast caused by the cellular objects in the light path. As the cells are generally semi-transparent and cell organelles are too small to be distinguished clearly with the wavelength of the white light source, the cell images in bright field data look somewhat like texture patches. Previously, several object recognition techniques have been tried to detect and count cells in bright field images (4, 5). Texture analysis methods where cell contours were extracted after initial segmentation were presented by Korzynska et al (6) while Long et al (7) have argued that it is possible to classify cultured cells between different cell types without fluorescent stains. However, in their paper, the living cells were segmented manually. In addition, the cell population types are very different in their morphology and texture when compared to primary cells. When working with living primary cells such as macrophages in a large screening context, sub-classification of cell types may not be as simple because of the complex biology or purification method of the primary cells when compared to cultured ones. Proving automation

3

on cultured cells is a first step in this paper. We have moved one step further in both automation of analysis of primary cells and the classification of healthy and infected cells based on a number morphological, texture and frequency features calculated from bright field data.

Literature review shows that there is not much work done on automatic analysis of amplitude contrast bright field cell data. We did not find any work on amplitude contrast bright field images of primary macrophage cells in a high throughput screening context. HTS screening is particularly lacking in suitable methods for direct cell density measurements rather than biochemical fluorescence which do not allow analysis of individual cells across a distribution. In this context, the present work provides a novel method to address these needs. In this paper, we propose a sequential set of techniques that will convert the bright field cytological image data into relatively high contrast images facilitating the use of region growing techniques to identify and label individual cells automatically. We then calculate hundreds of unbiased features for each segmented cell. A weighted linear feature vector is then used to separate the healthy cells from the infected ones. In large high throughput, high content screening studies, especially on live cells, tuning acquisition parameters during image acquisition is not a reasonable option resulting in high image variation than in fixed cells. This work shows successful analysis which should only improve in fixed sample studies. In this work, we have added a LED light source to a high throughput screening instrument, to collect bright field images. The fluorescent images of the same macrophages were then also obtained with a live and dead cell stains for an unbiased validation of our method.

4

Material and Methods

Monocyte Derived Macrophage Culture: Negatively selected, aphered CD14 positive monocytes were purchased from Biological Specialty Corporation (BSC). Monocytes were shipped on dry ice and immediately placed in liquid nitrogen vapor phase for storage until ready for use. For plating, cells were thawed quickly at 37C and washed in complete media. Complete media consisted of RPMI 1640 (Hyclone) supplemented with 10% FBS (Sigma), 2mM L-Glutamine (Invitrogen), 40U/mL GM-CSF (R&D Biosystems), and 100 U/mL M-CSF (R&D Biosystems). Total cell count and viability was determined using Trypan blue (Sigma) and a hemacytometer. Cell density was adjusted to $7x10^5$ cells/mL and 50uL of cell suspension (35,000 cells) was added to each well of a 384 well Cell Carrier plate (Perkin Elmer). Plated cells were incubated at 37^oC in humidified atmosphere with 5% CO_2 for 24 hours. Typical monocyte adherence was 25-30%. Media was replaced 24 hours and 96 hours post plating. Cells were differentiated for a total of 7 days. Day 7 MDMs were infected with *F. tularensis ΔblaB*::GFP, an Ft LVS mutant that stably expresses green fluorescent protein (GFP) integrated onto the chromosome under the control of a constitutive promoter, at a multiplicity of infection (MOI, bacteria/ cell) of 100. MDMs and bacteria were incubated at 37^oC/5% CO_2 for 2 hours to allow for phagocytosis after which unincorporated bacteria were removed by washing 3 times with HBSS. Infected cells were incubated at 37°C and 5% CO_2 for 50 to 72 hours (32 wells of data for every two hour infection period). These cells that are infected for such a long period where microbial multiplication is visible can be generally considered as fully infected population. Another well-plate contains uninfected cells as control and fixed using formaldehyde at the same time as infection plate.

5

Image Acquisition: An Opera confocal imaging system (www.cellularImaging.com/products/Opera) was modified with a simple blue LED flashlight hung over the 384 well-plate to act as bright field light source as shown in the photograph in Fig. 1. There was no automation in place for fast switching of the bright field light source to collect bright field and fluorescent image simultaneously. First, bright field images were collected from all of the image fields in a well and then the same image fields were reacquired in fluorescent mode for comparison by scanning over the well a second time with the flash light source switched off. As the images within a well were acquired in a few minutes, the spatial changes in the cell locations between bright field imaging and fluorescent imaging can be considered negligible.

Images from fifty fields per well were collected in all infected and control wells. Each image field generally contains about thirty cells thereby giving over 5000 infected cells and 5000 control cells per infection time point for analysis. A 40X water immersed objective with NA=0.9 and collected transmitted light in the range of 400nm to 600 nm. The pixel intensities were binned by a factor of 2 during the camera level quantization resulting in images of 20X resolution with better signal to noise ratio.

Image Analysis: The general image analytics flow diagram to detect and classify the cells is shown in Fig. 2. Image analysis was divided into image enhancement, cell segmentation and control/infected classification. It is a common knowledge that amplitude contrast image does not match well with known edge based segmentation techniques (8, 9) and region growing techniques (10, 11). The main reason for the failure of the traditional segmentation technique

6

on the amplitude contrast data is that at higher resolution, the data is highly textured with both coarse and fine texture components which are indicative of light absorption by the cellular ultrastructures. We have resolved this issue by using gradient magnitude map of the bright field image to improve the object contrast prior to segmentation.

Image Enhancement: The bright field amplitude contrast images are characterized by low contrast with uneven and noisy background as shown in example Fig. 3. (12). Selinummi et al have used a novel z-projection technique where the images are acquired at different focal planes and projected on to a single plane to reduce the noise and enhance the contrast arguing that by using suitable projection methods one could decrease the background noise and enhance the contrast. We have chosen a simpler option, where, by increasing the camera exposure time we can increase the signal to noise ratio and obtain relatively smooth background and then if required, further improve the contrast digitally.

The first step in the image enhancement is to enhance the contrast of the cells in the bright field images. We propose to combine histogram equalized, diffused gradient magnitude map, and the original gray scale image followed by the grayscale rescaling to enhance the contrast. If I is the input image and g_{eq} is the histogram equalized gradient magnitude image then the contrast enhanced image $I_{contrast}$ is given by

$$I_{contrast} = \max(I) \cdot \frac{(I \oplus G) - \min(I \oplus G)}{\max(I \oplus G) - \min(I \oplus G)} \qquad ----- (1)$$

where, $G = \dfrac{1}{2\pi\varsigma^2} e^{-\frac{x^2+y^2}{2\varsigma^2}}$ is the Gaussian diffused image of the histogram equalized gradient

map. A standard deviation $\varsigma = 21$ is used for Gaussian diffusion.

Fig. 4(A) shows an example bright field image and Fig. 4(B) shows contrast enhanced image by the above method.

The second step is to eliminate background intensity variation. This is accomplished by reconstructing a background image as a very low frequency version of the contrast enhanced image. Fig 5(a) shows the background image reconstructed by low-pass filtering of the input image by large sigma. The background image is then subtracted from the smoothed contrast enhanced image. The negative values were clipped to zero and the result is rescaled. The resulting image is further diffused using Gaussian filter ($\varsigma = 3$) to facilitate binarization. Fig. 5(b) shows background suppressed and smoothed foreground image.

Image Segmentation: Segmentation of the individual macrophages in the low-resolution image is achieved by the following steps.

Step 1: The images are threshold to obtain a global mask with a clear separation of background and foreground pixels. We have implemented minimum error thresholding method (13) for this purpose. If we threshold the gray level data at a value T and model each of the two resulting pixel populations by a normal density $h(g|i,T)$ with parameters $\propto_i(T), \varsigma_i(T)$ and prior probability $P_i(T)$ then the minimum error threshold T can be calculated by minimizing the following criterion,

8

$$J(T) = 1 + 2 \oplus \{P_1(T)\log(\lceil_1(T)) + P_2(T)\log(\lceil_2(T))\} \square 2 \oplus \{P_1(T)\log(P_1(T)) + P_2(T)\log(P_2(T \quad ---(2)$$
$$))\}$$

Where,

$$P_g(T) = \square_{g=a}^{b} h(g)$$

$$\propto_g(T) = \frac{\square_{g=a}^{b} g \oplus h(g)}{P_g(T)} ; \quad \lceil_g(T) = \frac{\square_{g=a}^{b} (g \square \propto_g(T))^2 \oplus h(g)}{P_g(T)} \quad a = \begin{cases} 0 & i=1 \\ \lceil T+1 & i=2 \end{cases} ; \quad b = \begin{cases} T & i=1 \\ \lceil n & i=2 \end{cases}$$

√‾‾‾‾‾‾‾‾‾‾
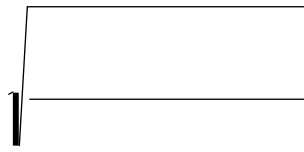
Step 2: The holes in the binary image were filled using a hole-filling algorithm. The negative binary image was labeled and all the objects in the negative image that are under certain size are considered as representing holes in the positive image and hence those pixels were restored to foreground intensity in the thresholded image. In our experiments all the objects in the negative image that are less than 50% of the normal cell size were considered as holes.

Fig. 6 shows result of thresholding by minimum error method.

Step 3: This step describes the segmentation of individual cells by region growing from automatically generated seeds. The binary image of cell clusters is converted to synthetic gray scale image using Euclidean distance transform (14). A peak search algorithm is used to mark the local peaks in the distance map. A peak is a pixel or a group of pixels whose distance is larger than or equal to the maximum distance value of the immediate neighborhood pixels. After all the local peaks were marked, a peak selection technique is used to eliminate most of

the multiple peaks within a single cell. The peaks were clustered into groups based on their proximity to each other. All the peaks that are within a certain distance from each other and the line connecting the two peaks does not pass through background are clustered into a same group. The peaks within the cluster are given a same label and marked indicating that they represent the same cell. The distance value of all the pixels in a marker (i.e. peaks with same marker label) is updated to the maximum distance value of the peaks with the same marker label. This step reduces some of the major cell fragmentation which is the main disadvantage of the watershed like method.

In the first iteration of the region growing, the markers that have maximum distance value in the whole image become active marker(s) and rest of the markers are considered passive for

that iteration of region growing. Active markers are grown into its neighboring pixels under a distance constraint. Let d_p where $p = 1,2..., n$ is the distance value of n pixels in the immediate neighborhood of the active marker m. Let d_m the distance value of the corresponding active marker. The active marker is grown into the pixel p_n in the neighborhood if $p_n \, \varepsilon \, (d_m \square 1)$. Growing the marker is implemented by labeling the unlabeled neighborhood pixels that satisfy the distance condition. The distance value of all the pixels within a grown region i.e. active marker, were then updated to the minimum distance value within the grown region i.e. if n_m is the set of pixels in the active marker 'm', then the updated distance $d_m = \min(d_p)$ where $p = 1,2,..., n_m$. In the next iteration, the passive markers whose distance value $d_m \, \varepsilon \, \min(d_{am})$ where $am = 1,..., \text{active markers}$, will become new additional active

markers. The region growing is stopped when all the foreground pixels in the binary image are assigned a corresponding marker label.

In spite of all the efforts to avoid multiple markers per cell, it is not possible to completely remove the cell fragmentation caused by noisy markers. Peak clustering process generally reduces the fragmentation caused by multiple peaks near the center of the cells which, otherwise, difficult to rectify. The noisy peaks near the cell boundary caused by less than perfect convex surface of the cells would still cause fragmentation of the cell although these fragments are quite small and can be detected by size and shape parameters. A post processing stage where these fragments and their parental cells were automatically detected and merged is thus necessary.

Step 4: To detect the fragments, we have calculated the depth of objects as the maximum distance value within the labeled objects and if this distance value is less than a certain preset distance, the object is considered as a fragment. If the radius of an average cell is r pixel units then the objects whose depth value are less than $\frac{r}{4}$ pixel units are considered as cell fragments. For each detected cell fragment, all the possible parental cells (touching neighbors that are not fragments themselves) were identified. The fragment is merged with a parental cell that has a longest shared boundary with the fragment. Fig. 7 shows intermediate results of segmentation process on one of the images. Fig. 8 shows the result of segmentation on several image samples with the boundary of segmented image overlaid on the original images. Note that the robustness of the filtering and the segmentation steps is such that there is no need of

retuning theses parameters for small changes in the image quality. Fig. 8 shows images from different wells with differing quality which are all segmented automatically without tuning any analysis parameters. Thus the process renders itself very well with high throughput screening requirements.

Feature Analysis: Bright field data of macrophages, unlike fluorescent imaging data, does not show any deterministic and/or functional feature(s) that can be used to determine infectivity of a cell. It is thus important to quantify all possible features from the cell image and provide the feature weights to design a suitable classifier. Most of the features are rotation and scale invariant in nature. Cell size, shape features and their derivatives depend on the scaling and hence are suitably normalized before classification. We have calculated over one-thousand features for each cell image. These features include basic size and shape descriptors (15), multi-scale features (16), invariant moment features (17), statistical texture features (18), Laws texture features (statistical texture features of Laws texture images) (19), differential features of the intensity surface (features of local gradient magnitude, local gradient orientation, Laplacian, Isophote, Flowline, Brightness, shape index, etc.) (20), frequency domain features, histogram features, distribution features (radial, angular, etc. of intensity distribution, gradient magnitude distribution), local binary pattern image features (21), local contrast pattern image features (22), cell boundary features, edge features and other heuristic features such as spottyness, Chi-square distance between histograms of pixel patches, between concentric circular areas within the cell, gray class distance, heuristic and problem specific features, etc.

Features were calculated for each cell and normalized to have a value between 0.0-to-1.0. If we assume that each feature is Normally distributed, Jain and Dubes method (23) of normalization where the normalized feature value $\tilde{f}_i = \dfrac{f_i \square \propto_i}{\Gamma_i}$ where f_i is the set of values of ith feature for the cells in training data, $(\propto_i, \Gamma_i)$ are the mean and standard deviation describing the distribution, has a probability of 68% to be in the range $\begin{bmatrix} \square 1 & 1 \end{bmatrix}$. Aksoy and Haralick (24) have proposed additional shift and rescaling as,

$$\tilde{f}_i = \frac{\dfrac{f_i \square \propto_i}{3\Gamma_i} + 1}{2} \quad ----(3)$$

which guarantees 99% of $\tilde{f}_i$ to be in the $\begin{bmatrix} 0 & 1 \end{bmatrix}$ range. We have truncated the out of range components to either 0 or 1.

Control/Infected Classification: Our classification problem is a simple one. We have to reduce the dimensionality of the feature vector in the first step and find a threshold to classify the objects into two classes. Although Principal Component Analysis (PCA) is the widely used dimensionality reduction method, it has two major disadvantages for using in the current application. It is inflexible for using different dissimilarity measures other than Euclidean distance and is quite poor in preserving the distances between the classes. The multidimensional scaling (MDS) introduced by Shepard (25) and later by Kruskal (26) overcomes the above disadvantages of PCA. The purpose of the MDS is to provide a structure to the sample distribution which attempts to satisfy dissimilarity matrix constructed from the sample

feature vectors. There are many commercial software and freeware available to use the MDS to analyze samples. We have used Primer software (www.primer-e.com) that was originally designed to do ecological sample analysis in marine biology (27) but can be extended in principle to the current problem.

Results

Fig. 9 shows the consistency within the data class. We have randomly sampled about 50 cells from control wells and viewed MDS ordination plots with Euclidean distance as the (dis)similarity measure. Similar experimentation was done on infected cell population. Several different sets of samples were taken in a bootstrap mode to test the data consistency within a class. Considering the quality of the imaging as well as uneven transition of monocytes to macrophages, we have found a small variance in the control cell feature space. The variance of the infected cell feature space was somewhat larger. This is because of two main reasons. First, there were a number of uninfected cells in the infected wells. Second, the infected cell sample was collected from wells that are infected between 50 hrs to 72 hrs thus showing possible additional phenotypic variation in feature space. Fig. 10 shows hierarchical clustering of cell samples randomly selected from the micro-plate wells of control cells and the infected cells. Fig. 11 shows the 2D MDS ordination plot of the cell samples based on Euclidean distance as dissimilarity value. All the 'I' indicate that the sample is drawn from the micro-plate well that was infected and 'C' indicates the sample is drawn from the well that was uninfected (control). We can use the dendogram of Fig. 11 to impose the cluster on the MDS configuration plot as shown by drawing possible regions belonging to the same cluster. Fig. 12 shows the MDS

15

ordination plot of over four hundred cells drawn from infected and control wells. Dissimilarity measures such as average distance between the cell sample features, binomial deviance and Euclidean distance are used. This is done to demonstrate that by using a different or problem specific dissimilarity measure, one can better classify the bright-field cell samples in the phenotypic feature space.

Discussion

In general, we have demonstrated that it is possible to use bright-field amplitude contrast image data to classify the cells into different phenotypic classes by using robust image analysis methods utilizing a large number of feature descriptors. In most primary screens in drug discovery projects, researchers are looking for a cell density, viability in correlation with some other basic cell features. This paper demonstrates that in such cases one need not have to use expensive dyes achieve the primary screen goals. This allows for rapid assay development and actually preserves spectral options for stains if increased dimensionality is desired. Therefore, this method provides important enhancement to improve screening cost, speed and accuracy.

The image analysis routines and feature descriptors are written in .NET framework using C# language making them widely accessible. By using techniques such as non-metric MDS for reducing the feature space dimensionality, one can choose the dissimilarity measure that more-or-less preserves the original distance between the inherent classes in the data. In the future work, we plan to use industrial standard phase contrast high throughput microscopy to find the biological stages as unbiased classes of data in a large scale infection modeling experiment.

References

1. Ambriz-Colin, F; Torres-Cisneros, M; Avina-Cervantes, J. G; Saavedra, Martinez, J. E; Debeir, O; Sanchez-Mondragon, J. J. Detection of biological cells in phase-contrast microscopy images. *IEEE Artificial Intelligence Conference* **2006,** *5,* pp. 68-77.

2. Hand, A. J; Sun, T; Barber, D. C; Hose, D. R. and MacNeil, S. Automated tracking of migrating cells in phase-contrast video microscopy sequences using image registration. *Journal of Microscopy* **2009,** *234*, pp. 62-79.

3. Debeir O; Ham P.V; Kiss R; Decaestecker C; tracking of migrating cells under phase contrast video microscopy with combined mean shift processes. *IEEE transactions on Medical Imaging*, **2005,** *24*, pp.697-711.

4. Tscherepanow M; Zollner F; Hillebrand M; Kummert F; Automatic segmentation of unstained living cells in bright field microscope images, P. Perner and O. Salvetti (Eds.): MDA 2008, LNAI 5108, Springer-Verlag Berlin Heidelberg 2008, pp. 158-172.

5. Tscherepanow M; Zollner F; Classification of segmented regions in bright field microscope images, Proceedings of 18th International conference on Pattern Recognition, ICPR-06, 2006, pp. 972-975.

6. Korzynska A; Strojny W; Hoppe A; Wertheim D; Hoser P; Segmentation of microscope images of living cells, *Pattern Analysis Applications*, **2007**, *10,* pp. 301-319.

7. Long X;; Cleveland W.L; Yao L. Y; Multiclass cell detection in bright field images of cell mixtures with ECOC probability estimation, *Image and Vision computing*, **2008**. *26*, pp. 578-591.

17

8. P. S. Umesh Adiga: Segmentation of volumetric tissue images using constrained active contour models, *Computer Methods and Programs in Biomedicine,* **2003***, 71*, pp.91-104.

9. Ray, N, Acton, ST, and Ley, K, "Tracking leukocytes in vivo with shape and size constrained active contours," *IEEE Trans. On Medical Imaging*, **2002,** *21*, pp. 1222-1235.

10. Adams R; Bischoff L; seeded region growing, *IEEE transactions on Pattern Analysis and Machine Intelligence*, **1994,** *6*, pp. 641-643.

11. P. S. Umesh Adiga, B. B. Chaudhuri: An efficient method based on watershed and rule-based merging for segmentation of 3-D histo-pathological images. *Pattern Recognition,* **2001,** *34*, pp.1449-1458.

12. Selinummi J; Ruusuvuori P; Podolsky I; Ozinsky A; Gold E; Yli-Harja O; Aderem A; Shmulevich I; Bright field microscopy as an alternative to whole cell fluorescence in automated analysis of macrophage images, *Plos One*, **2009**, *4*, e7497.

13. Kittler J; and Illingworth J; , Minimum error thresholding, *Pattern Recognition*, **1986**, 19, pp. 141-47.

14. G. Borgefors, Distance transformations in digital images. *Computer Vision, Graphics and image Processing,* **1986,** *34*, pp. 344–371.

15. Paul Yushkevich, Sarang Joshi, Stephen M. Pizer, John G. Csernansky and Lei E. Wang, Feature Selection for Shape-Based Classification of Biological Objects, *Information Processing in Medical Imaging Lecture Notes in Computer Science*, **2003**, *2732*, pp.114-125.

16. Laptev I; Lindberg T; Tracking of multi-state hand models using particle filtering and a hierarchy of multi-scale features, In: M. Kerkhove Eds. *Scale-space and morphology in computer vision*, **2001,** LNCS *2106*, pp. 63-74.

17. Mohamed Rizon, Haniza Yazid, Puteh Saad, Ali Yeon Md Shakaff, Abdul Rahman Saad,Mohd Rozailan Mamat, Sazali Yaacob, Hazri Desa and M. Karthigayan, Object Detection using Geometric Invariant Moment, *American Journal of Applied Sciences,* **2006,** *2*, pp. 1876-1878.

18. Haralick R.M; Statistical and structural approaches to texture, *Proc. IEEE*, **1979**, 67, pp. 786-804.

19. Celia Varela, Nico Karssemeijer and Pablo G. Tahoces; Classification of Breast Tumors on Digital Mammograms Using Laws' Texture Features , *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2001* Lecture Notes in Computer Science, **2001**, *2208*, pp.1391-1392.

20. Srinivas S Ravela, "On multi-scale differential features and their representations for image retrieval and recognition" (January 1, 2003). *Electronic Doctoral Dissertations for UMass Amherst.* http://scholarworks.umass.edu/dissertations/AAI3078716

21. Matti Pietikäinen, Image Analysis with Local Binary Patterns, *Image Analysis*: *Lecture Notes in Computer Science*, **2005**, *3540*, pp. 115-118.

22. M.K. Bashar and N. Ohnishi, Image Retrieval by Local Contrast Patterns and Color, G. Bebis et al. (Eds.): ISVC **2006**, LNCS *4292*, Springer-Verlag Berlin Heidelberg, pp. 136-145.

23. Jain A.K. and Dubes R.C., Algorithms for clustering data, Prentice Hall, NJ, 1988.

24. Selim Aksoy, Robert M. Haralick, Feature Normalization and Likelihood-Based Similarity Measures for Image Retrieval, *Pattern Recognition Letters,* **2001,** *22* pp.563-582.

25. Shepard, R. N; The analysis of proximities: Multidimensional scaling with an unknown distance function I. *Psychometrika,* **1962***, 27*, pp. 125-140.

26. Kruskal, J. B; Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika,* **1964,** *29*, pp. 1-27.

27. K. Robert Clarke, Paul J. Somerfield, Laura Airoldi and Richard M. Warwick, Exploring interactions by second-stage community analyses, *Journal of Experimental Marine Biology and Ecology*, **2006**, *338*, pp. 179-192.

Fig 1: Photograph of SIMPLE CONVERSION of Opera Imaging System to Obtain Bright-Field Image
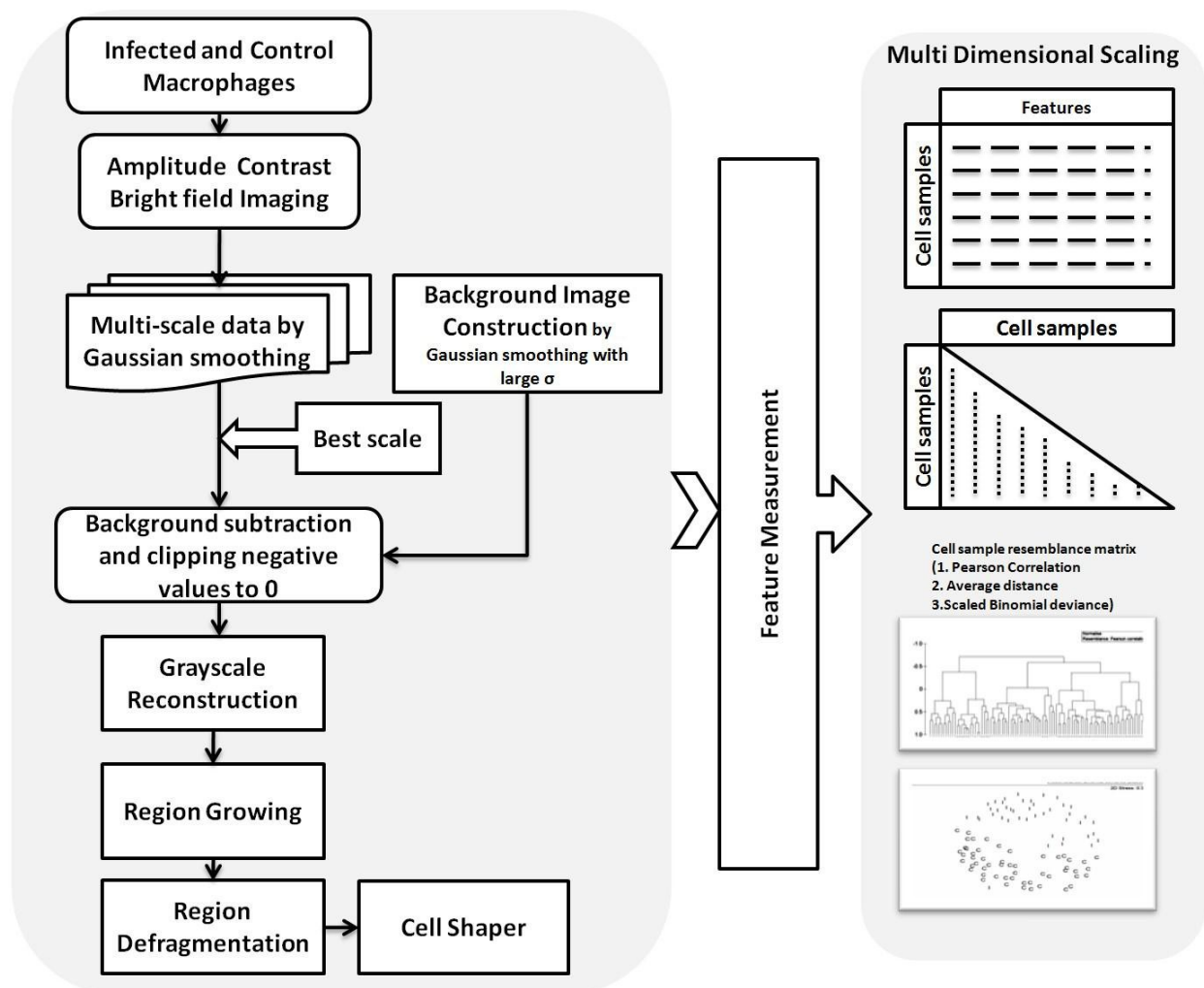
Fig 2: Control Flow Diagram of Bright-Field Cell Image Analysis Process
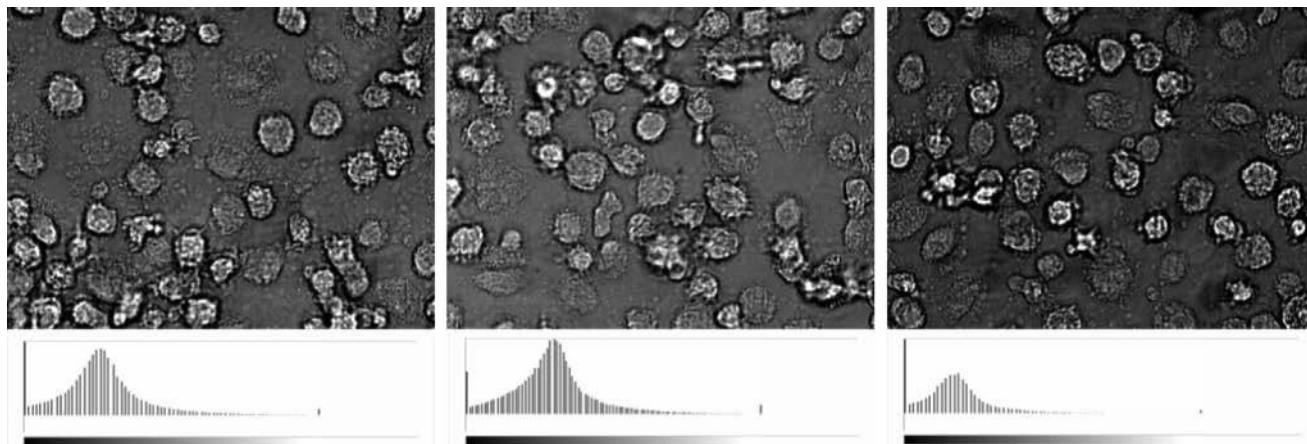
Fig 3: Set of Bright-Field Images taken from a Single Micro-Plate Well without Changing any

Data Acquisition Parameters. Corresponding histograms are also shown.
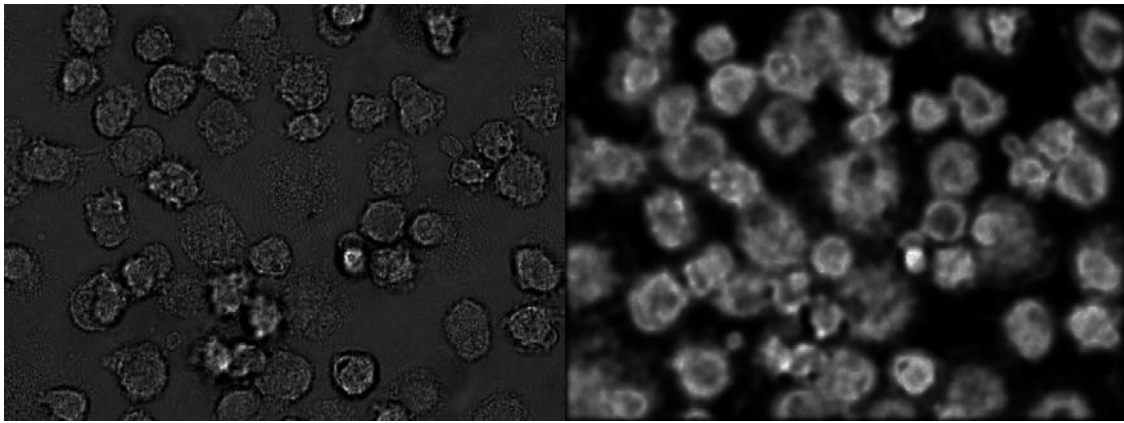
(A)   (B)

Fig. 4: Image Contrast Enhancement by Histogram Manipulation (A) Original Image (B) Contrast Enhanced Bright-Field Image
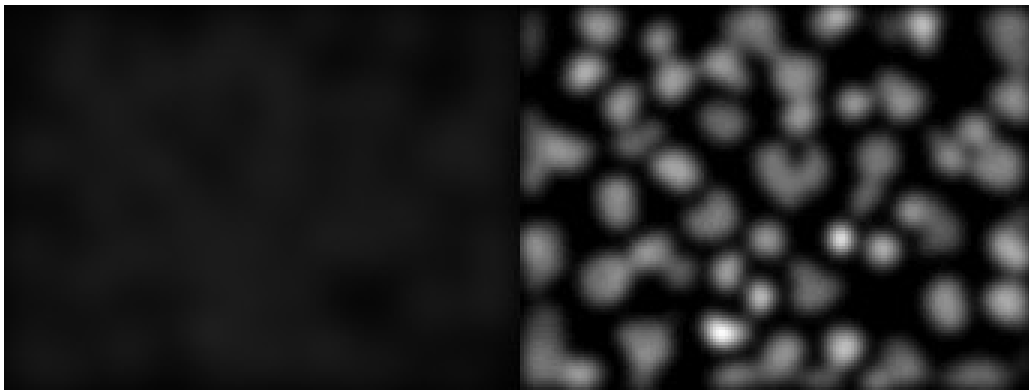
(A)   (B)

Fig. 5: Process of Background Intensity Elimination (A) Background Image (B) Background

Suppressed and Smoothed Foreground Image
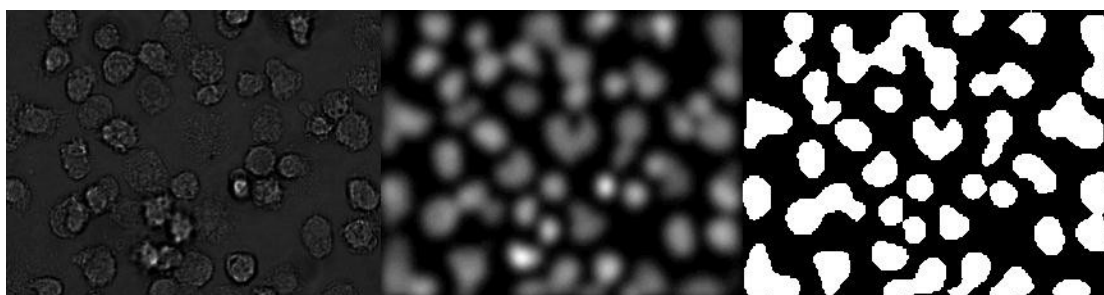
(A)   (B)   (C)

Fig. 6: (A) Original Image (B) Pre-Processed Image (C) Binary Mask By Minimum Error Thresholding
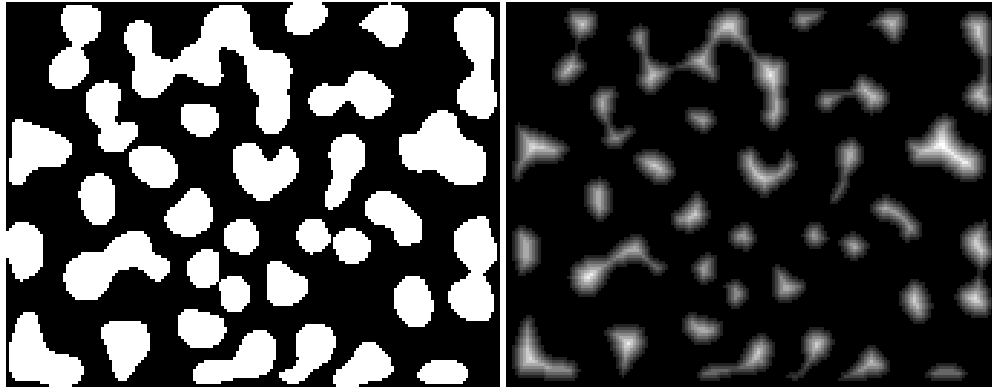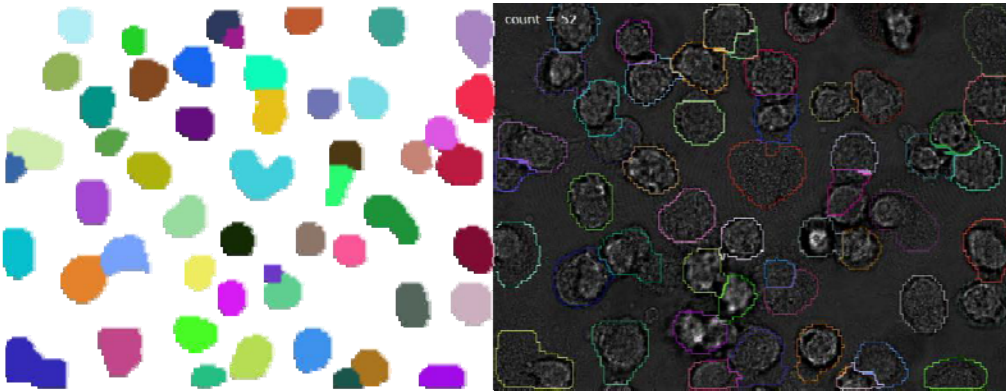
Fig. 7: Intermediate Results of Different Segmentation Steps (A) after minimum error thresholding, (B) reconstructed gray scale image using distance transform (C) segmentation result after artifact removal fragment merging and shaping the objects (D) Boundary of the masks in C is superposed on original data.
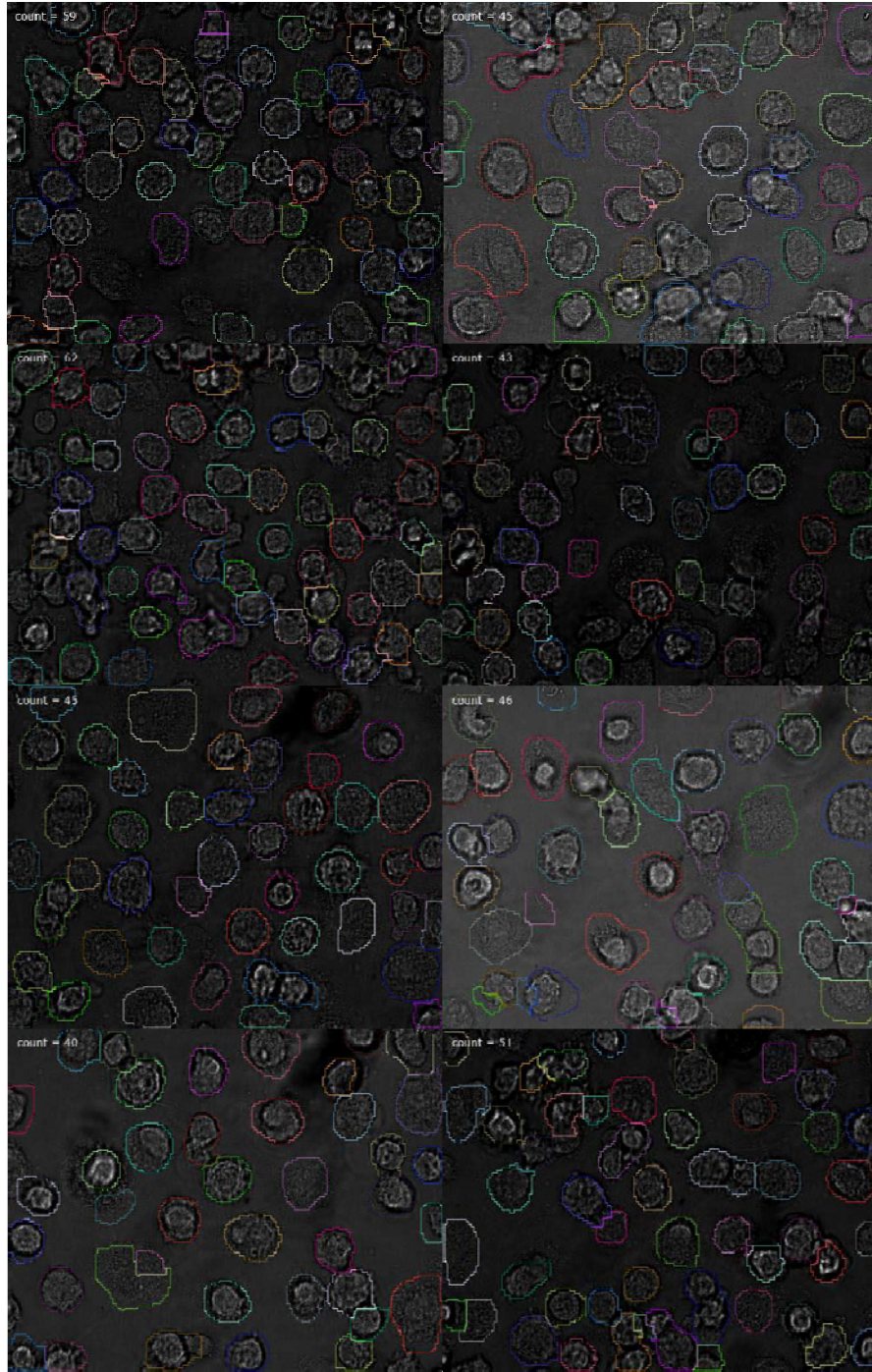
Fig 8: Examples of Amplitude Contrast Bright-Field Images of Macrophages from Micro Plate Wells that are Segmented by the above Described Method and the Mask Boundary is Superposed on the Original Gray Scale Image.
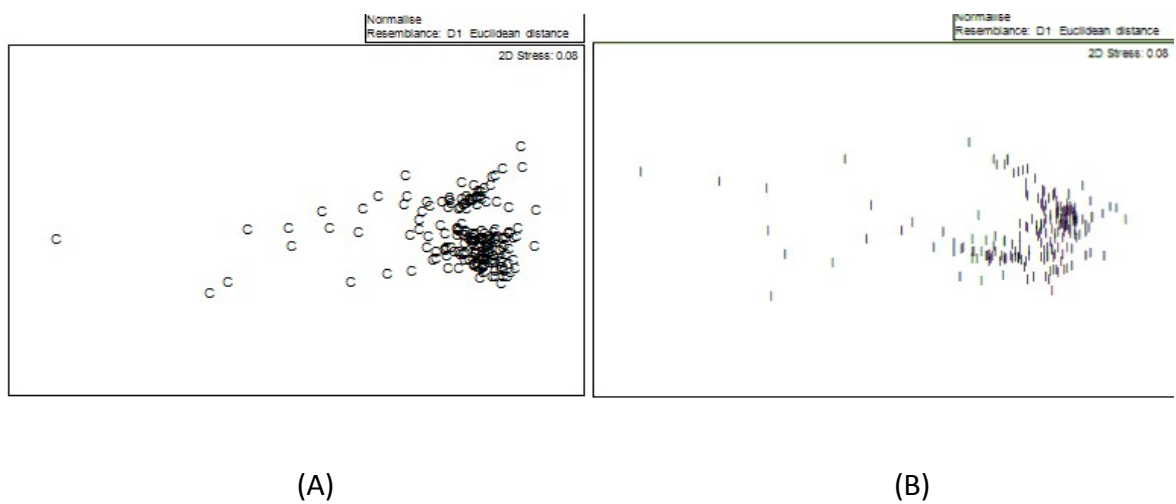
(A)   (B)

Fig. 9: A Non-Metric MDS Ordination Plot of (A) Samples Extracted from Control Wells (B) Samples from Infected Wells. Euclidean Distance is used as a Dissimilarity Measure.
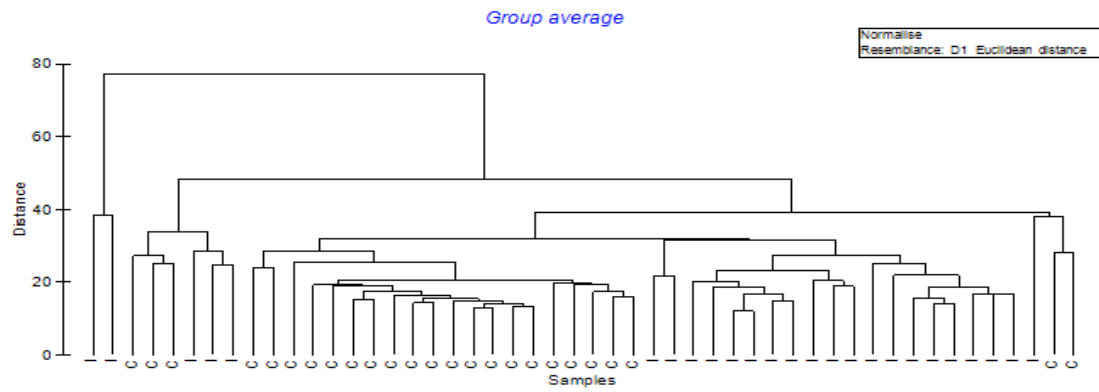
Fig 10: Hierarchical Clustering of the Cell Samples Based Pearson Correlation Coefficient Measure as a Similarity Value between Cell Samples. I: Infected, C: Control Cell Samples.
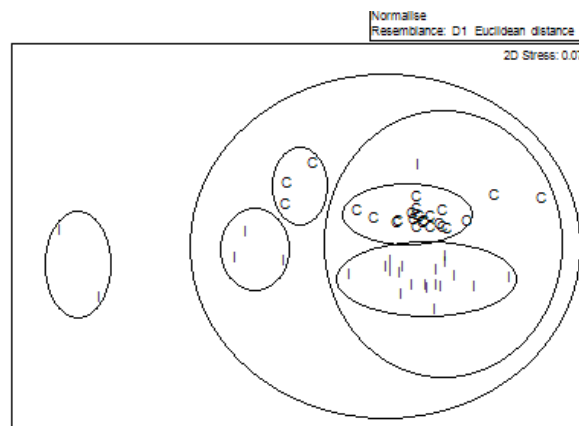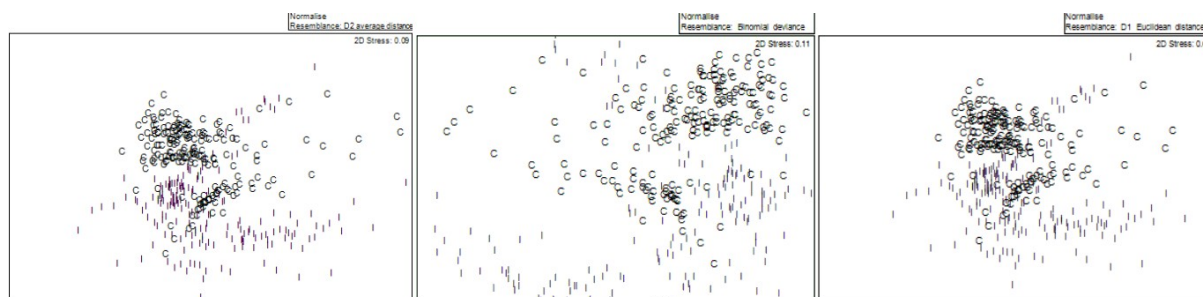
Fig. 11: Two dimensional MDS ordination plot of infected (or dead) and control sample points with Euclidean distance being used as (dis)similarity measure. I: Infected, C: Control cell samples.

(A)   (B)   (C)

Fig. 12(A): Two dimensional MDS ordination plot of infected and control sample points with average distance being used as dissimilarity measure (B) Two dimensional MDS ordination plot of infected and control sample points with Binomial deviance between sample point features as dissimilarity measure (C) Two dimensional MDS ordination plot of infected and control sample points with Euclidean distance as dissimilarity measure